Institute of Education

UCL

Survey methods research: new technologies

By Lisa Calderwood, Emily Gilbert, Emla Fitzsimons, David Bann, Praveetha Patalay, and Alice Sullivan

CENTRE FOR LONGITUDINAL STUDIES

UKRI Economic and Social Research Council

Summary

Project overview

This element of the project aimed to identify new technologies for data collection that could add scientific value to longitudinal studies including the CLS cohorts. We undertook an in–house literature review (by Emily Gilbert) and also employed 7 UCL PhD students to carry out thematic scoping reviews (whose remit covered data collection innovations and novel linkages related to scientific themes). The PhD students were employed as research assistants on short-term (8-week) contracts in May and June 2019 and supervised by senior academic staff within CLS.

New technologies

Technological change has transformed our daily lives in the last 20 years, and advances in technology create important opportunities for collecting data at scale. Smartphones are owned by two-thirds of adults in the UK and there has been explosion in the availability of data derived from people's daily interactions through the internet and elsewhere. CLS studies have used new technologies for data collection in MCS (age 14) which included a time-use diary which was mixed-mode using web, Smartphone app and paper, and in MCS (age 7 and age 14) surveys and BCS70 (age 46) surveys which used a portable activity monitor to collect direct measurements of physical activity. The challenge for CLS, and other longitudinal studies, is how to make best use of these new opportunities in the future to strengthen future research.

Literature review

Methods

The output covered here is a review of the methodological literature and evidence on the use of new technologies and innovative methods in data collection. The review looked at what major longitudinal studies have already done in terms of collecting data using new technologies and innovative methods, as well as innovation in other academic and market research settings. This covers the opportunities and challenges in the use of new technologies and wearables for both passive and active data collection, including cost, measurement issues and take-up rates.

Findings

- There are a large number of opportunities to enhance traditional survey data collected by large-scale longitudinal studies by collecting data through innovative techniques. There are also a number of methodological challenges surrounding innovative data collection, including willingness of participants to take part in non-standard methods of data collection, take up rates and representation, measurement issues, ethics, practicalities for large-scale longitudinal studies, and cost. More research is needed to understand potential biases introduced through these measures, particularly to understand who participates and who does not, to identify exactly what is being measured by different devices, and whether it can be considered a replacement for asking survey questions about certain behaviours, or whether these are complimentary measures.

Thematic scoping reviews

Methods

We decided at the outset of the project to take a thematic approach to scoping new data collection opportunities, and that as well as covering new technologies it was also important to scope possibilities for data enhancements through novel linkages. As a result, these reviews are broader than just new technologies for data collection.

Six scoping reviews were carried out in several key scientific areas. These reviews consisted of primarily desk-based research, including grey literature. They were conducted by UCL PhD students who were employed as research assistants on short-term (8-week) contracts in May and June 2019 and supervised by senior academic staff within CLS.

The key domains covered are: mental health, physical activity, cardiovascular health, diet, cognition, digital use, and social media. Each review discusses the opportunities and challenges around novel measurement.

Findings

- For measuring **mental health**, we scoped out available technology and feasibility of Apps, wearables and momentary assessment tools that have been used to-date, mainly in small scale studies. We discussed the capability of these approaches to measuring mental health and the considerations for scaling these up to population-based cohort studies.
    - We found a number of smartphone apps available that can collect both active (self-report) and passive (device usage information, and location) data, allowing for continuous monitoring of mental health.
    - Wearable technology has been employed for measuring mental health, including the use of technology to infer mental state from breathing rate, heart rate, and EEG sensors.
    - Ecological momentary assessment has also been widely used to assess anxiety and depression in particular.
    - Device standardisation was found to be a particular issue, particularly around the aggregation of data from different types of smartphones.
    - Wearable technology raises some issues in terms of the reliability of devices, as well as the obtrusiveness of some methods of measurement (such as the use of bands worn around the head).
    - User feedback was found to be a useful method of increasing engagement and cooperation in the use of new technologies to measure mental health.

- Wearable **physical activity** monitors continue to advance, and typically decrease in cost, yet yield data which are challenging to process and interpret; novel methods include combined wearable monitors and the use of ‘smart’ mobile phones which typically passively collect activity data for long periods of time.
    - The most promising approach to a combined wearable system is the combination of heartrate monitors with accelerometers, to improve accuracy of measurement. Combined systems such as this may be better at capturing activities that traditional accelerometers fail to capture, such as cycling or weightlifting.
    - However, combined wearables are more expensive than accelerometers alone, and may also increase participant burden if they involve the use of multiple devices or more invasive devices.
    - Smartphones also provide an opportunity for physical activity measurement, through the use of their in-built sensors. This may be a more cost-effective approach, and reduce participant burden.

- Indicators of **cardiovascular health** have historically been obtained by standard blood pressure and blood draws; novel alternatives include the use of dried blood spots, and tests of cardiorespiratory fitness.
    - Dried blood spots are a relatively inexpensive method of obtaining biomarkers, as well as being less invasive than venepuncture. Analytical methods are also improving, increasing the range of biomarkers detectable through dried blood spots.
    - Devices for measuring blood pressure have been developed that do not use a cuff, allowing for continuous monitoring outside of a medical setting. However, these devices are currently costly, making them less viable for population-based research at this stage.
    - Cardiorespiratory fitness can be measured in-field using shuttle run tests or submaximal step tests. Both of these methods have reasonably strong correlations with direct measures of oxygen uptake measured in laboratory settings.

- Multiple complementary measures of **diet** continue to advance, including means of self-reported intake (paper and electronic) and objectively assessed biomarkers. We discuss these methods, their potential for inclusion in cohorts, and describe a series of available resources to help guide selection of future dietary intake measures.
    - Mobile phones have been used to collect dietary information, due to their widespread use and the potential for real-time data collection. There are a wide number of commercial apps that track diet. Barcode scanning can also be incorporated through a smartphone, to allow collection of information related to purchased or consumed food.
    - Web-based dietary diaries are also reasonably common, with a large number of web based tool available.
    - Camera and tape recorder technologies have been used to measure dietary intake through either visual or verbal records of consumption, along with plate waste, which is then analysed to calculate nutritional intake
    - Wearables have been designed to record food intake through cameras, microphones and other sensory methods. The major benefit of wearables, or other approaches that use cameras and sensors, is that there is less reliance on respondent memory.

- Continuous development and increasingly widespread use of technology facilitates the administration of **cognitive** tests in population-based studies. Novel technologies harnessed for enriched datasets on cognition comprise mobile (smartphone and tablet) applications, web-based cognitive testing, wearable technology, smart home systems, and non-invasive neural interface technology.
    - Smartphones have been used for short-term memory tasks with respondents, enabling collection of this information from participants whilst they are in their homes, reducing burden.
    - Web-based approaches have also been used, again reducing burden by allowing completion in-home. Some batteries of cognition tasks have already been validated for web administration.
    - Wearable technologies can be used to assess physical states, psychological states, social interactions, and environmental context. Several studies have utilised wearables to cognitively monitor and assist the elderly. This data can often come from participants' own devices, reducing burden and cost.
    - Smart home systems have been developed to assist people experiencing cognitive decline. Whilst not designed to measure cognition directly, there is

potential to infer cognitive function by observing changes in activities of daily living.
    - Non-invasive neural interface technology can also be used to assess cognition. EEGs and fNIRs have both been used to assess participants whilst they are completing cognitive assessments. Wireless EEG has the potential to be used to assess cognition on large-scale studies.

- **Social media** is increasingly used for research in a range of areas including mental health, politics and social capital. Linking social media to survey data provides the opportunity to validate measures derived from social media with validated survey measures. However, there are a number of methodological challenges to be considered including consent, ethics and data processing.
    - Social media data obtained from well-known groups (i.e. longitudinal cohorts) can enhance researchers' understanding of human behaviour and, for example, how a person's mental health changes over time.
    - University of Bristol is currently developing a framework for linking and sharing social media data for high-resolution longitudinal measurement of mental health across CLOSER cohorts.
    - Few studies that have used social media data have cross-verified their findings with external datasets or captured demographic information. Linking social media data with survey data would add value in terms of representivity, validity and reliability.
    - Informed consent is a challenge in this space - a key element of asking for informed consent is full transparency over the type of data being linked, the purposes of collection, data security and participants' rights to withdraw their consent.

- It is important for large-scale studies to measure the increasing **digitisation** of lives. Self-reported measures of online behaviours are not able to capture the full range of these activities. There are a number of technological solutions that passively measure people's online activities, though issues around ethics and data security need to be considered.
    - Browser plug-ins utilise a small piece of software installed by the user which tracks how much time is spent on websites visited and which websites are visited. Desktop time tracking software looks beyond websites and gathers information on what's happening with the device operating system (which apps are being used, for how long). In a similar vein, smartphone tracking apps can be installed on a smartphone to track which apps the user is using and for how long.
    - Cursor tracking can be used to collect the position of the cursor on the screen.
    - Session replay has the ability to replay a visitor's journey to a website, seeing what the user sees and how they interact with the website.
    - Browser fingerprinting is a technique which collects information about browser type and version, operating system, plugins, language, screen resolution
    - Tracking cookies can be used to track users' web browsing habits across multiple websites.
    - Deep packet inspection (DPI) is a type of data processing that inspects in detail the data being sent over a computer network (which can then be use to assess which websites the user has visited and how they have interacted with them).
    - It should be borne in mind that technology use is not limited to one device or a specific activity, and user activity is constantly changing, which should drive decisions about which method is best to capture the required information.

Next steps and recommendations

The findings from this work package will be used by CLS to inform the design of future data collection in the cohort studies. We will also draw on other information, including from the CLS event on large-scale surveys and technology innovation event in June 2019. We will identify and select a set of novel measures in key areas such as screen time, diet, mental health to be trialled in the future for use in the cohort studies. In the forthcoming BCS70 Age 50 Survey, we are intending to pilot collecting data from an app which captures objectively measured smartphone screen time and a second app, a smartphone game (SeaHeroQuest) which will measure cognitive skills. These reports and findings will also be of much wider interest and utility for other large-scale surveys incorporating new technologies for data collection. We summarise the broader implications of these findings below.

There are a large number of opportunities to enhance traditional survey data collected by large-scale longitudinal studies by collecting data using new technologies. The possibilities for more finely granulated and real-time measurement are attractive for research, and it seems feasible that new technologies could be used to collect data in a wide range of topic areas. There is also evidence to suggest that different populations can also be persuaded to participate in data collection exercises using new technologies – from young children to the elderly, lots of different groups have participated in these activities.

Smartphones appear to be a front-running candidate for the collection of data on numerous topics. As well as being able to collect active data, through surveys or momentary assessments, the plethora of sensors built in to these devices offers huge opportunity for data collection. Studies focusing on travel and movement, physical activity, health behaviours, consumer behaviour, and digitisation of lives, to name a few examples, could also be built around smartphone data collection. The fact that smartphone ownership is widespread make them a potentially cost-effective solution for collecting data, as well as possibly reducing burden among participants.

Wearables are also becoming a more common solution for capturing objective data that can be used in a variety of different analyses. There are huge number of commercially available devices which can measure things such as heart rate, physical activity and blood pressure, useful metrics in studies across a broad range of disciplines. Wearables are also able to capture contextual information, such as photographs or audio recordings. These can be used to understand many facets of daily life, as well as validate information captured using alternative methods. In addition, data from participants' own wearable devices may be utilised to answer some research questions in a cost-effective way.

However, more research is needed to understand willingness and compliance. Bias can be introduced through the use of new technologies and innovative data collection methods: who chooses to participate in innovative data collection exercises, and how are they different to the population of interest? We also need more work to identify exactly what is being measured by different devices, and how to aggregate data from different types of device, and devices running different software.

Given the rapidly changing nature of this sphere, the ethics, practicalities and costs associated with innovative data collection are ever-changing. Keeping up with developments will be paramount in ensuring large-scale longitudinal studies can utilise the opportunities offered by new technologies and innovative data collection methods to provide the research community with relevant, high-quality, valuable data.